WEST GRID

compute canada | calcul canada

Updates about local HPC resources

UofM-Spring-Workshop 2021

April 21st-22nd, 2021

Grigory Shamov

University of Manitoba

University of Manitoba

Introductions:

Grigory Shamov (Team Lead, Research computing support)
Dr. Ali Kerrache (HPC Specialist, Research Computing support)

Housekeeping: Working with GoogleMeet.

Video call link: https://meet.google.com/hxz-nsef-exe
Or dial: (CA) +1 226-314-8840 PIN: 269 720 725#

- Please mute your mic if not talking
- Please don’t share your screen yet
- Chat is available to ask questions during presentation.

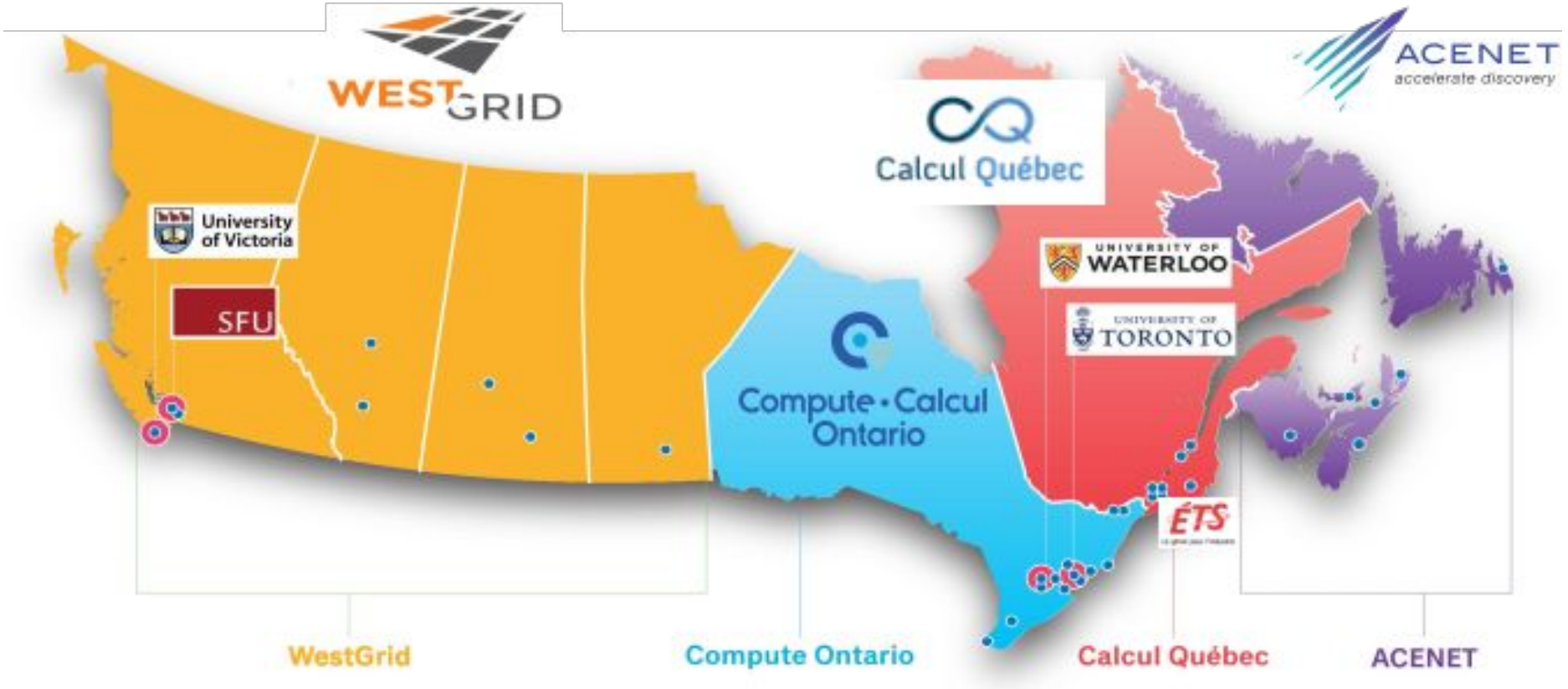
Thanks!

University of Manitoba

Outline

- ★ Available resources (2:00 - 2:15 PM, April 21):
  - Compute Canada updates, National resources
  - Grex (UofM HPC resource); **hardware updates** and **RAC-2021**
- ★ Basics of using HPC clusters (2:15 - 3:00 PM, April 21):
  - Linux shell (Terminal)
  - Connecting to a cluster, Transfer of data
  - Submit and monitor batch jobs with SLURM
- ★ Break at 3PM.
- ★ More in-depth discussion: (3:05PM - 4:00 PM, April 21)
  - SLURM best practices
  - HPC software
- ★ Q/A, office hours, finishing up in-depth topics on Apr 22.

(the old) Compute Canada



Compute Canada Systems

System	Cores	GPUs	Storage	Notes
Cedar	94K	1352	29 PB	HPC machine, has P100; V100 Volta GPUs
Graham	42K	520	19 PB	HPC machine, P100; V100; T4 GPUs
Beluga	28K	688	27 PB	HPC machine, has V100 GPUs
Niagara/ Mist	80K	216	2 PB	Large parallel jobs; [4 NVIDIA V100-32GB]
Arbutus	16K	108	17.3 PB	OpenStack Cloud, virtualized V100 GPUs
GP cloud	*	*	*	Cloud partitions are available on HPC systems for special purposes.

https://docs.computecanada.ca/wiki/Compute_Canada_Documentation

ComputeCanada Services

- National ARC helpdesk support@computecanada.ca
  - Support of the CC HPC systems and private clouds
- National ARC software stack CC-CVMFS
- Various research portals and projects like MagicCastle, JupyterHubs, Galaxy, etc; Globus data transfer and sharing platform, NextCloud.
- Outreach and training events
  - Run by consortia
  - Westgrid Summer School is about to start on April 27!
- Grant consultations, User contributed systems support
- Resource allocation process (RAC)

New DRI organization

NDRIO to replace ComputeCanada in 2022

- ComputeCanada ceases to exist
- NDRIO takes over existing CC Systems and Support
- NDRIO adds RDM (like FRDR) and Research Software to its portfolio

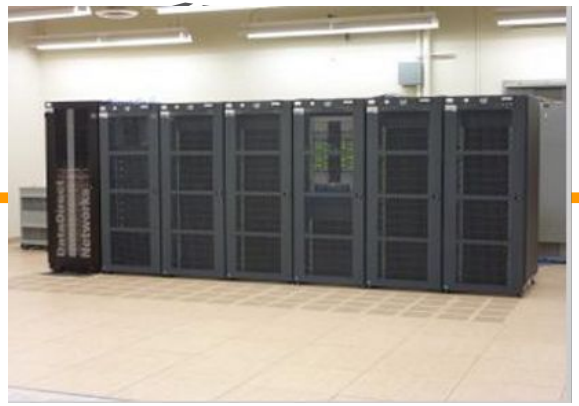
Consultations / Input wanted from researchers:
DRI Needs assessment Virtual Town halls

- https://engagedri.ca/canadian-digital-research-infrastructure-needs-assessment/virtual-town-halls
- Registration is open untill April 30
- Sessions on Awareness, Governance, Operations and Support on May 4, 5, 6, 7 correspondingly
- Everyone (PIs, researchers, librarians, Administrators, ARC support staff) is welcome to attend

The local HPC resource, Grex

- Grex is a formerly National machine from 2010, which we inherited after its defunding
- Provides a traditional HPC system capacity for local users
  - Cost efficient , high utilization, managed software stacks, etc.
  - Used by many ; user base is more or less the same with UM users of ComputeCanada
- A long term supplementary resources for local users that cannot it get elsewhere
- Helps to ramp up local users to usage of national HPC systems
  - Similar user experience (CCDB, SLURM, software) standardization on same technologies
- Make Grex into a “community cluster” (by adding contributed systems)
  - Accept and manage user-contributed hardware in a standard HPC way
  - Allow for better resource sharing and TCO reductions for the PIs
  - General trend for mid-size HPC systems that do not receive National funding

Grex, old and new hardware

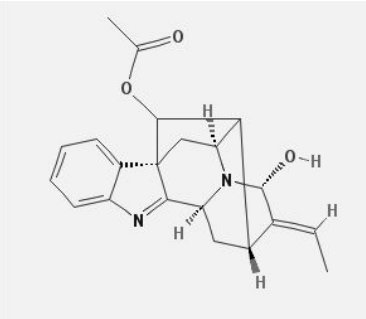


Original Grex, SGI nodes racked in HPCC

New compute nodes from Lenovo!
Thanks to IST funding

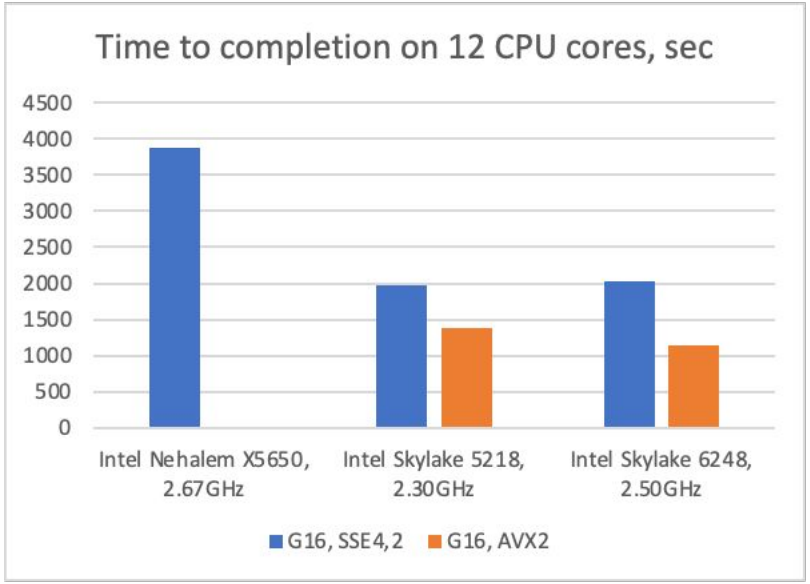
Node type	# Total /Online	CPUs Total /Online	Status, running or planned	Plans?
SGI Grex compute, 12 Intel X5650, 48Gb RAM	316 / 233	3792 / 2796	since Nov. 2010	Decommission in Oct 2021 ?
SGI Grex login	2 (public)		since Nov. 2010	Decommission in Oct 2021 ?
Intel High-mem 40 Intel 6248 CPU, 384Gb	12 / 11	480 / 440	since Sept. 2020	Production use
GPU nodes, 4xV100 GPU 32 Intel 5218 CPU, 192Gb	2 / 2	64 CPUs, 8 GPUs	since March 2020	Production use
Intel Low-mem nodes 52 Intel 6230R CPU, 96GB	**42 / 0**	**2184 / 0**	Delivered	Start in May 2021
new 6230R login nodes	**3/0**	**156 / 0**	Delivered	Start in May 2021

A benchmark for new CPUs



Vomilinene, $C_{21}H_{22}N_2O_3$
on 12 CPUs, 32Gb RAM
Gaussian 16, opt B3LYP/6-31G(d)

Partition	CPU model	CPU freq, GHz	Streaming instruction set
compute	**Intel Nehalem X5650**	2.67	SSE4.2
gpu	**Intel Skylake 5218**	2.30 (Turbo)	AVX512
skylake	**Intel Skylake 6248**	2.50	2x AVX512



- **Nehalem has SSE4.2, while Skylake got AVX512**
- **Largest benefit for codes that can use AVX512!!**

University of Manitoba

User contributed nodes

Origin of PI	Node kind	#	CPUs, GPUs Total	Status	Plans?
Chemistry	Intel E5v2, 2.2 GHz, 32Gb RAM, Chemistry	4	80 CPUs	since Aug. 2014	Keeping them
Physics	GPU nodes, 4xV100 GPU 16GB 32 Intel 5218 CPU, 192Gb	3	96 CPUs, 12 GPUs	since Nov. 2019	New nodes, production use
Comp. Science	GPU node, 16xV100 GPUs, Intel 6248R CPU, 1.5TB RAM	1	48 CPUs, 16 GPUs	since Feb. 2021	New server, production use

University of Manitoba

Storage

Type	Storage FS	Capacity, quota	Status	Plans?
/home	NFS, SAS disks, DDN	5 TB , 30Gb/user	Works since 2010	Decommission Jun 2021
/global/scratch	Lustre, SATA disks, DDN	110 TB, 1TB/user	2010 - 2017	-
/global/scratch	Lustre, SATA disks, Seagate	418 TB, 4TB/user	Works since 2017	in production
/home	NFS, NVME disks, SuperMicro	20TB, 80GB/user	Testing	Start Jun 2021

Managing Grex resources

With the new hardware and contrib hardware, Grex becomes a very heterogeneous system!

- We had to introduce hardware “partitions”,
  - Scheduling becomes less efficient
  - Users must know the resources!

- We’d need a decision on what to allocate
  - Old nodes will be decommissioned and thus not allocatable

- Call a new RAC round to update the resource allocations and gather your input.
  - The local Resource Allocation Call is out
  - deadline **May 5, 2021**

Local Resource Allocation Call

Long time passed since the 2019-2020 call. The new local RAC call is out. As usual, we are trying to minimize the hassle.

What do we allocate for this yer?

- Only new CPU cores (**2.6K,** Intel AVX512)
- Storage is allocated as before (on /global/scratch)
- GPUs are first come-first serve, counted as CPUs
  - *We ask to indicate demand for GPUs on the applications*
- Old nodes will be decommissioned mid-way of the RAC year
  - *We ask to indicate demand for old CPUs on the applications*
- Contributed systems are not allocated but open for preemptible jobs
  - *The jobs will have to be accounted differently from jobs ran on the allocatable HW*

University of Manitoba

Local RAC

RAS (the defaut) and RAC streams

- RAS if estimated CY is no more than 30 and storage no more than 10TB/group
- RAC if either of compute core years or storage needs to be larger
- After evaluation of all applications, the asks might be scaled to fit into allocatable resources

We do appreciate progress reports!

- Publications using Grex, achievements enabled by it
- We need to justify our funding by Research Manitoba and it helps
- We need to know what is going on on Grex for our capacity planning at UManitoba